```
           Copyright   GenCore version 5.1.7
           Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on:    February 14, 2006, 17:53:18 ; Search time 45 Seconds
                                (without alignments)
                                1821.704 Million cell updates/sec

Title:          US-10-612-012-2
Perfect score:  852
Sequence:       1 MSSVSPIQIPSRLPLLLTHB.........AAFDGGFTVKTRPGLLNSKL 852

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       283416 seqs, 96216763 residues

Word size :     10

Total number of hits satisfying chosen parameters:    50

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :    PIR_80:*
              1: pirl:*
              2: pir2:*
              3: pir3:*
              4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2.0	17	774	2	B84031	ATP-dependent prot
2	2.0	17	795	1	A70322	endopeptidase La (
3	1.9	16	784	2	AC0383	endopeptidase La (
4	1.9	16	875	2	T09142	endopeptidase La h
5	1.8	15	779	1	B42375	endopeptidase La (
6	1.6	14	798	2	G83420	Lon proteinase PA1
7	1.6	14	827	1	A36895	endopeptidase La (
8	1.6	14	848	2	C82712	ATP-dependent seri
9	1.5	13	773	2	T27382	hypothetical prote
10	1.5	13	778	2	E97224	ATP-dependent Lon
11	1.5	13	786	2	A96956	ATP-dependent prot
12	1.5	13	799	1	F83549	probable ATP-depen
13	1.5	13	817	2	A49844	endopeptidase La (
14	1.5	13	821	2	B75530	ATP-dependent prot
15	1.5	13	885	1	T04321	endopeptidase La (
16	1.5	13	964	1	T04325	probable ATP-depen
17	1.4	12	784	1	S47270	endopeptidase La (
18	1.4	12	784	2	SUECLA	endopeptidase La*(
19	1.4	12	784	2	AB0558	Lon protease (impo
20	1.4	12	784	2	B90690	endopeptidase La [
21	1.4	12	787	2	A72230	endopeptidase La (
22	1.4	12	799	2	A85541	hypothetical prote
23	1.4	12	805	2	AH2731	ATP-dependent prot
24	1.4	12	805	2	G97512	ATP-dependent prot
25	1.4	12	810	2	JC6045	endopeptidase La (
26	1.4	12	820	2	H81106	ATP-dependent prot
27	1.4	12	820	2	H81908	probable endopepti
28	1.4	12	881	1	B71316	endopeptidase La (
29	1.4	12	1067	1	S62421	endopeptidase La h
30	1.4	12	1096	1	T01765	endopeptidase La-1
31	1.4	12	1133	1	S43938	endopeptidase La-1
32	1.3	11	779	2	E97778	endopeptidase La (
33	1.3	11	784	1	A71704	endopeptidase La (
34	1.3	11	786	1	G82141	ATP-dependent LA p
35	1.3	11	795	1	D64226	endopeptidase La (
36	1.3	11	803	1	A64070	endopeptidase La (
37	1.3	11	813	1	D70176	endopeptidase La (
38	1.3	11	819	1	B72128	endopeptidase La (
39	1.3	11	819	1	C71527	endopeptidase La (
40	1.3	11	819	2	B86494	endopeptidase La (
41	1.3	11	819	2	E81681	Lon ATP-dependent
42	1.2	10	774	1	I40421	proteinase, Lon fa
43	1.2	10	777	2	F84985	endopeptidase La (
44	1.2	10	791	2	D82901	ATP-dependent prot
45	1.2	10	795	1	S73830	endopeptidase La (

ALIGNMENTS

```
RESULT 1
B84031
ATP-dependent proteinase La 1 (lon) (class III heat-shock protein) lonA [imported] - Ba
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: B84031
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B84031
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-774 <STO>
A;Cross-references: UNIPROT:Q9K8P6; UNIPARC:UPI0000C40AD; GB:AP001517; GB:BA000004; NII
A;Experimental source: strain C-125
C;Genetics:
A;Gene: lonA
C;Superfamily: ATP-dependent Lon protease

Query Match       2.0%;   Score 17;   DB 2;   Length 774;
Best Local Similarity   100.0%;   Pred. No. 4.6e-08;
Matches   17;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   453 GDPAALLEVLDPEQNH 469
         ||||||||||||||||
Db   433 GDPAALLEVLDPEQNH 449


RESULT 2
A70322
endopeptidase La (EC 3.4.21.53) - Aquifex aeolicus
N;Alternate names: ATP-dependent proteinase lon; ATP-dependent serine proteinase La
N;Contains: adenosinetriphosphatase (EC 3.6.1.3)
C;Species: Aquifex aeolicus
C;Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 09-Jul-2004
C;Accession: A70322
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:9816566; PMID:9537320
A;Accession: A70322
A;Molecule type: DNA
A;Residues: 1-795 <AQF>
A;Cross-references: UNIPROT:O66605; UNIPARC:UPI00000562D8; GB:AE000680; NID:g2982948; PI
C;Comment: This allosteric enzyme catalyzes the hydrolysis of large proteins in the pres
C;Genetics:
A;Gene: lon
C;Superfamily: ATP-dependent Lon protease
```

```
                GenCore version 5.1.7
       Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model
Run on: February 14, 2006, 17:49:34 ; Search time 189 Seconds
                                         (without alignments)
                         1980.692 Million cell updates/sec

Title:          US-10-612-012-2
Perfect score:  852
Sequence:       1 MSSVSPIQIPSRLPLLLTHE.........AAFDGGFTVKTRPGLLNSKL 852

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0
Searched:       2443163 seqs, 439378781 residues
Word size :     10
Total number of hits satisfying chosen parameters:   149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  A_Geneseq_21:*
            1: geneseqp1980s:*
            2: geneseqp1990s:*
            3: geneseqp2000s:*
            4: geneseqp2001s:*
            5: geneseqp2002s:*
            6: geneseqp2003as:*
            7: geneseqp2003bs:*
            8: geneseqp2004s:*
            9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	852	100.0	852	5	ABG32363	Abg32363 Human ATP
2	852	100.0	852	7	ADM56266	Adm56266 Human ATP
3	852	100.0	852	9	ADV85873	Adv85873 Human pro
4	815	95.7	852	4	AAB47562	Aab47562 Protease
5	815	95.7	852	7	ADC10040	Adc10040 Human NOV
6	714	83.8	852	4	AAM93785	Aam93785 Human pol
7	714	83.8	852	5	ABP69443	Abp69443 Human pol
8	714	83.8	852	8	ADL31771	Adl31771 Human pro
9	693	81.3	693	7	ADJ70423	Adj70423 Human hea
10	581	68.2	581	4	AAB95332	Aab95332 Human pro
11	581	68.2	581	4	AAB99179	Aab99179 Human ATP
12	470	55.2	571	4	AAB94288	Aab94288 Human pro
13	432	50.7	432	4	AAB97167	Aab97167 Human, ATP
14	432	50.7	432	4	AAE11085	Aae11085 Human ATP
15	421	49.4	433	6	ABU11611	Abu11611 Human MDD
16	379	44.5	581	8	ADR08963	Adr08963 Human pro
17	331	38.8	423	4	AAU17375	Aau17375 Novel sig
18	331	38.8	423	7	ADB94083	Adb94083 Human nov
19	215	25.2	228	4	AAU17371	Aau17371 Novel sig
20	215	25.2	228	7	ADB94079	Adb94079 Human nov
21	195	22.9	391	5	ABB89248	Abb89248 Human pol
22	160	18.8	160	5	ADP55489	Adp55489 Human PRO
23	118	13.8	118	4	AAB01658	Aab01658 Human gen
24	85	10.0	85	8	ABO57485	Abo57485 Human gen
25	78	9.2	78	4	AAM14167	Aam14167 Peptide #
26	78	9.2	78	4	ABB33114	Abb33114 Peptide #
27	78	9.2	78	4	AAM26575	Aam26575 Peptide #
28	78	9.2	78	4	ABB27940	Abb27940 Human pep
29	78	9.2	78	4	ABB18580	Abb18580 Protein #
30	78	9.2	78	4	ABG47963	Abg47963 Human liv
31	78	9.2	78	5	AAM01905	Aam01905 Peptide #
32	78	9.2	78	8	ABG35946	Abg35946 Human pep
33	50	5.9	50	5	ABO57622	Abo57622 Human pep
34	32	3.8	102	5	ABP07459	Abp07459 Human gen
35	31	3.6	101	4	AAM91575	Aam91575 Human imm
36	21	2.5	82	4	AAU32043	Aau32043 Novel hum
37	17	2.0	774	8	ADS28353	Ads28353 Bacterial
38	17	2.0	786	8	ADN26262	Adn26262 Bacterial
39	17	2.0	795	8	ADN17367	Adn17367 Bacterial
40	16	1.9	767	8	ADN27141	Adn27141 Bacterial
41	16	1.9	777	8	ADN26487	Adn26487 Bacterial
42	16	1.9	784	6	ABU50220	Abu50220 Protein e
43	16	1.9	785	6	ABM66958	Abm66958 Photorhab
44	16	1.9	795	8	ADS42793	Ads42793 Bacterial
45	16	1.9	892	8	ADN20874	Adn20874 Bacterial

ALIGNMENTS

```
RESULT 1
ABG32363
ID   ABG32363 standard; protein; 852 AA.
XX
AC   ABG32363;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human ATP-dependent protease.
XX
KW   Human; enzyme; ATP-dependent protease; developmental disorder; cancer;
KW   retinoblastoma; melanotic melanoma; endometrial adenocarcinoma;
KW   ovarian adenocarcinoma; schizophrenia.
XX
OS   Homo sapiens.
XX
PN   US2002081704-A1.
XX
PD   27-JUN-2002.
XX
PF   21-DEC-2000; 2000US-00741150.
XX
PR   22-NOV-2000; 2000US-0252410P.
XX
PA   (GUEG/) GUGBLER K.
PA   (WEBS/) WEBSTER M.
PA   (YANC/) YAN C.
PA   (SHAO/) SHAO W.
PA   (KETC/) KETCHUM K A.
PA   (DFRA/) DI FRANCESCO V.
PA   (BEAS/) BEASLEY E M.
XX
PI   Guegler K, Webster M, Yan C, Shao W, Ketchum KA, Di Francesco V;
PI   Beasley EM;
XX
DR   WPI; 2002-635469/68.
DR   N-PSDB; ABK90887, ABK90888.
XX
PT   New human ATP-dependent peptides and encoding nucleic acids, useful for
PT   diagnosing, preventing and/or treating disorders like cancer, e.g.
PT   retinoblastomas, melanomas, endometrial and ovarian adenocarcinomas, and
PT   schizophrenia.
XX
PS   Claim 1; Fig 2; 119pp; English.
XX
CC   The invention relates to a new isolated ATP-dependent protease (I) and
CC   the nucleic acid encoding it. The peptides are useful in identifying
```

```
                    GenCore version 5.1.7
          Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on:     February 14, 2006, 17:58:02 ; Search time 169 Seconds
                                         (without alignments)
                                         2106.453 Million cell updates/sec

Title:           US-10-612-012-2
Perfect score:   852
Sequence:        1 MSSVSPIQIPSRLPLLLTHE........AAFDGGFTVKTRPGLLNSKL 852

Scoring table:   OLIGO
Gapop 60.0 , Gapext 60.0

Searched:    1867569 seqs, 41789326 residues

Word size :  10

Total number of hits satisfying chosen parameters:    124

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   Published Applications AA_Main:*
             1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
             2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
             3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
             4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
             5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
             6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	852	100.0	852	3	US-09-741-150-2	Sequence 2, Appli
2	852	100.0	852	5	US-10-612-012-2	Sequence 2, Appli
3	815	95.7	852	5	US-10-239-032-4	Sequence 4, Appli
4	693	81.3	693	4	US-10-408-765A-2229	Sequence 2229, Ap
5	432	50.7	432	4	US-10-156-239-44	Sequence 44, Appl
6	331	38.8	423	3	US-09-764-868-940	Sequence 940, App
7	215	25.2	228	3	US-09-764-868-936	Sequence 936, App
8	195	22.9	391	4	US-10-264-237-1624	Sequence 1624, Ap
9	85	10.0	85	4	US-10-029-386-31119	Sequence 31119, A
10	78	9.2	78	3	US-09-864-761-33878	Sequence 33878, A
11	50	5.9	50	4	US-10-029-386-31256	Sequence 31256, A
12	50	2.0	774	4	US-10-369-493-17386	Sequence 17386, A
13	17	2.0	786	4	US-10-369-493-8915	Sequence 8915, Ap
14	17	2.0	795	4	US-10-369-493-20	Sequence 20, Appl
15	17	2.0	670	4	US-10-424-599-183947	Sequence 183947,
16	16	1.9	767	4	US-10-369-493-9794	Sequence 9794, Ap
17	16	1.9	777	4	US-10-369-493-9140	Sequence 9140, Ap
18	16	1.9	784	4	US-10-282-122A-78144	Sequence 78144, A
19	16	1.9	795	4	US-10-369-493-21223	Sequence 21223, A
20	16	1.9	892	4	US-10-369-493-3527	Sequence 3527, Ap
21	16	1.8	280	4	US-10-425-114-43912	Sequence 43912, A
22	15	1.8	376	4	US-10-424-599-198007	Sequence 198007,
23	15	1.8	766	4	US-10-369-493-16141	Sequence 16141, A
24	15	1.8	768	4	US-10-282-122A-45835	Sequence 45835, A
25	15	1.8	769	4	US-10-369-493-16622	Sequence 16622, A
26	15	1.8	777	4	US-10-369-493-373	Sequence 373, App
27	15	1.8	784	4	US-10-282-122A-68623	Sequence 68623, A
28	15	1.8	788	4	US-10-369-493-11178	Sequence 11178, A
29	15	1.8	790	4	US-10-369-493-15757	Sequence 15757, A
30	15	1.8	794	4	US-10-282-122A-45766	Sequence 45766, A
31	15	1.8	811	4	US-10-156-761-10503	Sequence 10503, A
32	15	1.8	844	4	US-10-369-493-15389	Sequence 15389, A
33	14	1.6	783	4	US-10-369-493-9459	Sequence 9459, Ap
34	14	1.6	784	4	US-10-369-493-9196	Sequence 9196, Ap
35	14	1.6	798	4	US-10-369-493-13739	Sequence 13739, A
36	14	1.6	798	4	US-10-282-122A-66328	Sequence 66328, A
37	14	1.6	798	4	US-10-282-122A-67928	Sequence 67928, A
38	14	1.6	798	4	US-10-282-122A-69728	Sequence 69728, A
39	14	1.6	827	4	US-10-369-493-19432	Sequence 19432, A
40	14	1.6	837	4	US-10-369-493-4083	Sequence 4083, Ap
41	14	1.6	848	4	US-10-369-493-17569	Sequence 17569, A
42	14	1.6	954	4	US-10-369-493-3120	Sequence 3120, Ap
43	13	1.5	170	4	US-10-424-599-237202	Sequence 237202,
44	13	1.5	304	4	US-10-767-701-38439	Sequence 38439, A
45	13	1.5	388	4	US-10-282-122A-62236	Sequence 62236, A

ALIGNMENTS

```
RESULT 1
US-09-741-150-2
; Sequence 2, Application US/09741150
; Publication No. US20020081704A1
; GENERAL INFORMATION:
; APPLICANT: GUBGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Human
US-09-741-150-2

Query Match             100.0%;  Score 852;  DB 3;  Length 852;
Best Local Similarity   100.0%;
Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAAHNLQLVRSRLLKGTSLQSTILGVI 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAAHNLQLVRSRLLKGTSLQSTILGVI 60

Qy   61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120

Qy  121 IAEVEQLDRLEEFPNTCKMRBELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 IAEVEQLDRLEEFPNTCKMRBELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180

Qy  181 EALPDILTSIIRTSNKEKQLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 EALPDILTSIIRTSNKEKQLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240

Qy  241 RVIAIRPIRRITHISGTLEDEDEDENDDIVMLEKKIRTSSMPEQAHKVCVKBIKRLKKGM 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 RVIAIRPIRRITHISGTLEDEDEDENDDIVMLEKKIRTSSMPEQAHKVCVKBIKRLKKGM 300

Qy  301 PQSMPEYALTRNYLELMVBLPWNKSTTDRLDIRAARILLDNDHYAMEKLKGRVLEYLAVR 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  301 PQSMPEYALTRNYLELMVBLPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAVR 360

Qy  361 QLKQNLKGPILCFVGPPGVGKTSVGRSVAKTLGREFHRIALGGVCDQSDIRGHRRTYVGS 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                    GenCore version 5.1.7
             Copyright (c) 1993 - 2006  Biocceleration Ltd.

OM protein - protein search, using sw model

Run on:        February 14, 2006, 17:58:56 ; Search time 17 Seconds
                                            (without alignments)
                                            657.671 Million cell updates/sec

Title:         US-10-612-012-2
Perfect score: 852
Sequence:      1 MSSVSPIQIPSRLPLLTHE.........AAFDGGFTVKTRPGLLNSKL 852

Scoring table: OLIGO
               Gapop 60.0 , Gapext 60.0

Searched:      97014 seqs, 13122538 residues

Word size :    10

Total number of hits satisfying chosen parameters:      4

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :     Published Applications AA_New:*
               1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
               2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
               3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
               4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
               5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
               6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
               7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
               8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

```
Result          %    Query
No.    Score Match Length DB  ID                          Description
----   ----- ----- ------ --  --                          -----------
  1     13   1.5    830   7   US-11-098-686-10393          Sequence 10393, A
  2     12   1.4    804   7   US-11-070-080-18             Sequence 18, Appl
  3     12   1.4    820   6   US-10-467-657-4910           Sequence 4910, Ap
  4     11   1.3    817   7   US-11-098-686-11041          Sequence 11041, A
```

ALIGNMENTS

```
RESULT 1
US-11-098-686-10393
; Sequence 10393, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
;  APPLICANT: Kapur, Vivek and Gebhart, Connie J.
;  TITLE OF INVENTION:  NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
;  FILE REFERENCE: 09531-128001
;  CURRENT APPLICATION NUMBER: US/11/098,686
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: PCT/US03/31318
;  PRIOR FILING DATE: 2003-10-01
;  PRIOR APPLICATION NUMBER: US 60/416,395
;  PRIOR FILING DATE: 2002-10-04
;  NUMBER OF SEQ ID NOS: 11433
```

```
;  SEQ ID NO 10393
;  LENGTH: 830
;  TYPE: PRT
;  ORGANISM: Lawsonia intracellularis
US-11-098-686-10393

Query Match             1.5%;  Score 13;  DB 7;  Length 830;
Best Local Similarity   100.0%;  Pred. No. 0.00012;
Matches   13;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        685 LTGQLGDVMKESA 697
              |||||||||||||
Db        667 LTGQLGDVMKESA 679


RESULT 2
US-11-070-080-18
; Sequence 18, Application US/11070080
; Publication No.US20050287625A1
; GENERAL INFORMATION:
;  APPLICANT: Ye, Rick W.
;  APPLICANT: Miller, Edward S.
;  TITLE OF INVENTION: PROCESS FOR EXPRESSION OF FOREIGN GENES IN METHANE METABOLIZING
;  TITLE OF INVENTION:  BACTERIA THROUGH CHROMOSOMAL INTEGRATION
;  FILE REFERENCE: CL-2443 US NA
;  CURRENT APPLICATION NUMBER: US/11/070,080
;  CURRENT FILING DATE:  2005-03-02
;  PRIOR APPLICATION NUMBER: US 60/550385
;  PRIOR FILING DATE: 2004-03-05
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.3
;  SEQ ID NO 18
;  LENGTH: 804
;  TYPE: PRT
;  ORGANISM: Methylomonas sp. 16a
US-11-070-080-18

Query Match             1.4%;  Score 12;  DB 7;  Length 804;
Best Local Similarity   100.0%;  Pred. No. 0.0012;
Matches   12;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        458 ALLEVLDPEQNH 469
              ||||||||||||
Db        440 ALLEVLDPEQNH 451


RESULT 3
US-10-467-657-4910
; Sequence 4910, Application US/10467657
; Publication No. US2005026058lA1
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SpA
;  APPLICANT: FONTANA Maria Rita
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
;  SEQ ID NO 4910
;  LENGTH: 820
;  TYPE: PRT
;  ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4910

Query Match             1.4%;  Score 12;  DB 6;  Length 820;
Best Local Similarity   100.0%;  Pred. No. 0.0012;
```

```
              GenCore version 5.1.7
        Copyright  (c) 1993 - 2006  Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on:    February 14, 2006, 17:49:51 ; Search time 238 Seconds
                                (without alignments)
                           2525.673 Million cell updates/sec

Title:          US-10-612-012-2
Perfect score:  852
Sequence:       1 MSSVSPIQIPSRLPLLTHB.........AAPDGGFTVKTRPGLLNSKL 852

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       2166443 seqs, 705528306 residues

Word size :     10

Total number of hits satisfying chosen parameters:      265

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :    UniProt_05.80:*
          1: uniprot_sprot:*
          2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                            SUMMARIES
```

Result No.	Score	Query Match	Length	DB	ID	Description
1	815	95.7	852	2	Q86WA8_HUMAN	Q86wa8 homo sapien
2	715	83.9	715	2	Q8N3B9_HUMAN	Q8n3b9 homo sapien
3	714	83.8	852	2	Q8NCE9_HUMAN	Q8nce9 homo sapien
4	581	68.2	581	2	Q96K43_HUMAN	Q96k43 homo sapien
5	273	32.0	852	2	Q5R6M5_PONPY	Q5r6m5 pongo pygma
6	204	23.9	204	2	Q9BU35_HUMAN	Q9bu35 homo sapien
7	202	23.7	228	2	Q62MF9_HUMAN	Q62mf9 homo sapien
8	147	17.3	806	2	Q6TXI3_RAT	Q6txi3 rattus norv
9	147	17.3	852	2	Q9DBN5_MOUSE	Q9dbn5 m mus muscu
10	85	10.0	160	2	Q9D1A6_MOUSE	Q9d1a6 mus musculu
11	85	10.0	432	2	Q8BK80_MOUSE	Q8bk80 mus musculu
12	79	9.3	840	2	Q5PQY6_BRARE	Q5pqy6 brachydanio
13	69	8.1	114	2	Q8NF29_HUMAN	Q8nf29 homo sapien
14	30	3.5	874	2	Q4T681_TETNG	Q4t681 tetraodon n
15	21	2.5	222	2	Q9XKK8_LITER	Q9xkk8 lithospermu
16	18	2.1	489	2	Q6AJ30_DESPS	Q6aj30 desulfotale
17	18	2.1	774	2	Q6AK61_DESPS	Q6ak61 desulfotale
18	17	2.0	768	2	Q74EN9_GEOSL	Q74en9 geobacter s
19	17	2.0	774	2	Q9K8F6_BACHD	Q9k8f6 bacillus ha
20	17	2.0	795	1	LON_AQUAE	O66605 aquifex aeo
21	17	2.0	795	2	Q9JCX1_THETH	Q9jcx1 thermus the
22	17	2.0	795	2	Q72KS4_THET2	Q72ks4 thermus the
23	17	2.0	795	2	Q5SK78_THET8	Q5sk78 thermus the
24	17	2.0	817	2	Q83DJ0_COXBU	Q83dj0 coxiella bu
25	17	2.0	819	2	Q747S2_GEOSL	Q747s2 geobacter s
26	17	2.0	826	2	Q7NBA5_MYCGA	Q7nba5 mycoplasma
27	16	1.9	784	2	Q66DT2_YERPS	Q66dt2 yersinia ps
28	16	1.9	784	2	Q8ZC67_YERPE	Q8zc67 yersinia pe
29	16	1.9	784	2	Q7N0L5_PHOLL	Q7n0l5 photorhabdu
30	16	1.9	793	2	Q6D825_ERWCT	Q6d825 erwinia car
31	16	1.9	802	2	Q8D154_YERPE	Q8d154 yersinia pe
32	16	1.9	812	2	Q4NQW6_9DELT	Q4nqw6 anaeromyxob
33	16	1.9	843	2	Q4NPW6_9DELT	Q4npw6 anaeromyxob
34	16	1.9	875	2	LONH1_SPIOL	O04979 spinacia ol
35	16	1.9	937	2	Q7SA85_NEUCR	Q7sa85 neurospora
36	16	1.8	1147	2	Q6BJJ8_DEBHA	Q6bjj8 debaryomyce
37	15	1.8	557	2	Q8GT60_ARATH	Q8gt60 arabidopsis
38	15	1.8	702	2	Q725X1_DESVH	Q725x1 desulfovibr
39	15	1.8	773	2	Q4MS75_BACCE	Q4ms75 bacillus ce
40	15	1.8	773	2	Q722V6_BACC1	Q722v6 bacillus ce
41	15	1.8	773	2	Q81LC1_BACAN	Q81lc1 bacillus an
42	15	1.8	774	2	Q5QXP0_IDILO	Q5qxp0 idiomarina
43	15	1.8	776	2	Q6HSS4_BACAN	Q6hss4 bacillus an
44	15	1.8	776	2	Q817Q4_BACCR	Q817q4 bacillus ce
45	15	1.8	776	2	Q6HD56_BACHK	Q6hd56 bacillus th

```
                            ALIGNMENTS

RESULT 1
Q86WA8_HUMAN
ID   Q86WA8_HUMAN PRELIMINARY;      PRT;    852 AA.
AC   Q86WA8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Peroxisomal lon protease.
GN   Name=LONP;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   De Walque S., Van Veldhoven P.P.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RG   NIH MGC Project;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ548761; CAD68987.1; -; mRNA.
DR   EMBL; BC093912; AAH93912.1; -; mRNA.
DR   EMBL; BC093910; AAH93910.1; -; mRNA.
     HSSP; P08177; IRR9.
DR   MEROPS; S16.006; -.
DR   Ensembl; ENSG00000102910; Homo sapiens.
DR   GO; GO:0005524; F:ATP binding; IBA.
```

OM protein - protein search, using sw model

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: February 14, 2006, 17:57:16 ; Search time 49 Seconds
              (without alignments)
              1437.544 Million cell updates/sec

Title: US-10-612-012-2
Perfect score: 852
Sequence: 1 MSSVSPIQIPSRLPLLLTHE..........AAFDGGFVKTRPGLLNSKL 852.

Scoring table: OLIGO
               Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size :   10

Total number of hits satisfying chosen parameters: 16

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   Issued_Patents_AA:*
               1: /cgn2_6/ptcdata/1/iaa/5_COMB.pep:*
               2: /cgn2_6/ptcdata/1/iaa/6_COMB.pep:*
               3: /cgn2_6/ptcdata/1/iaa/H_COMB.pep:*
               4: /cgn2_6/ptcdata/1/iaa/PCTUS_COMB.pep:*
               5: /cgn2_6/ptcdata/1/iaa/RE_COMB.pep:*
               6: /cgn2_6/ptcdata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	852	100.0	852	2	US-09-741-150-2	Sequence 2, Appli
2	852	100.0	852	2	US-10-160-187-2	Sequence 2, Appli
3	15	1.8	790	2	US-09-543-681A-6059	Sequence 6059, Ap
4	14	1.6	810	2	US-09-902-540-11337	Sequence 11337, A
5	14	1.6	820	2	US-09-252-991A-30204	Sequence 30204, A
6	13	1.5	809	2	US-09-252-991A-29280	Sequence 29280, A
7	13	1.5	884	2	US-09-741-150-4	Sequence 4, Appli
8	13	1.5	884	2	US-10-160-187-4	Sequence 4, Appli
9	12	1.4	783	2	US-09-248-796A-16773	Sequence 16773, A
10	12	1.4	845	2	US-09-489-039A-13590	Sequence 13590, A
11	11	1.3	440	2	US-09-198-452A-44	Sequence 44, Appl
12	11	1.3	825	2	US-09-438-185A-29	Sequence 29, Appli
13	10	1.2	157	2	US-09-861-451A-8	Sequence 8, Appli
14	10	1.2	424	2	US-09-198-452A-45	Sequence 45, Appli
15	10	1.2	822	2	US-09-328-352-5754	Sequence 5754, Ap
16	10	1.2	832	2	US-09-540-236-3056	Sequence 3056, Ap

ALIGNMENTS

RESULT 1
US-09-741-150-2
; Sequence 2, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUBGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Human
US-09-741-150-2

Query Match   100.0%; Score 852; DB 2; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAHNLQVVRSRLLKGTSLQSTILGVI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MSSVSPIQIPSRLPLLLTHEGVLLPGSTMRTSVDSAHNLQVVRSRLLKGTSLQSTILGVI 60

Qy   61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 PNTPDPASDAQDLPPLHRIGTAALAVQVVGSNWPKPHYTLLITGLCRFQIVQVLKEKPYP 120

Qy  121 IAEVEQLDRLEEFPNTCKMRBELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 IAEVEQLDRLEEFPNTCKMRBELGELSEQFYKYAVQLVEMLDMSVPAVAKLRRLLDSLPR 180

Qy  181 EALPDILTSIIRTSNKEKLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 EALPDILTSIIRTSNKEKLQILDAVSLEERFKMTIPLLVRQIEGLKLLQKTRKPKQDDDK 240

Qy  241 RVIAIRPIRRITHISGTLEDEDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKKM 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 RVIAIRPIRRITHISGTLEDEDEDEDNDDIVMLEKKIRTSSMPEQAHKVCVKEIKRLKKM 300

Qy  301 PQSMPBYALTRNYLBLMVELPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAVR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 PQSMPEYALTRNYLBLMVELPWNKSTTDRLDIRAARILLDNDHYAMEKLKKRVLEYLAVR 360

Qy  361 QLKQNLKGPILCFVGPPGVGKTSVGRSVAKTLGREFHRIALGGVCDQSDIRGHRRTYVGS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 QLKQNLKGPILCFVGPPGVGKTSVGRSVAKTLGREFHRIALGGVCDQSDIRGHRRTYVGS 420

Qy  421 MPGRIINGLKTVGVNNPVFLLDEVDKLGKSLQGDPAAALLEVLDPEQNHNFTDHYLNVAF 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 MPGRIINGLKTVGVNNPVFLLDEVDKLGKSLQGDPAAALLEVLDPEQNHNFTDHYLNVAF 480

Qy  481 DLSQVLFIATANTTATIPAALLDRMEIIQVPGYTQBEKIBIAHRHLIPKQLEQHGLTPQQ 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 DLSQVLFIATANTTATIPAALLDRMEIIQVPGYTQBEKIBIAHRHLIPKQLEQHGLTPQQ 540

Qy  541 IQIPQVTTLDIITRYTRBAGVRSLDRKLGAICRAVAVKVABGQHKEAKLDRSDVTEREGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 IQIPQVTTLDIITRYTRBAGVRSLDRKLGAICRAVAVKVABGQHKEAKLDRSDVTEREGC 600

Qy  601 REHILEDEKPESISDTTDLALPEMPILIDFHALKDILGPPMYEMEVSQRLSQPGVAIGL 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 REHILEDEKPESISDTTDLALPEMPILIDFHALKDILGPPMYEMEVSQRLSQPGVAIGL 660

Qy  661 AWTPLGGEIMFVEASRMDGEGQLTLTGQLGDGVMKESAHLAISWLRSNAKKYQLTNAFGSP 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AWTPLGGEIMFVEASRMDGEGQLTLTGQLGDGVMKESAHLAISWLRSNAKKYQLTNAFGSP 720

Qy  721 DLLDNTDIHLHFPAGAVTKDGPSAGVTIVTCLASLFSGRLVRSDVAMTGBITLRGLVLPV 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 DLLDNTDIHLHFPAGAVTKDGPSAGVTIVTCLASLFSGRLVRSDVAMTGBITLRGLVLPV 780

Qy  781 GGIKDKVLAAHRAGLKQVIIPRRNEKDLEGIPGNVRQDLSFVTASCLDBVLNAAPDGGFT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GGIKDKVLAAHRAGLKQVIIPRRNEKDLEGIPGNVRQDLSFVTASCLDBVLNAAPDGGFT 840
```